Supplementary Figures 1 - 11

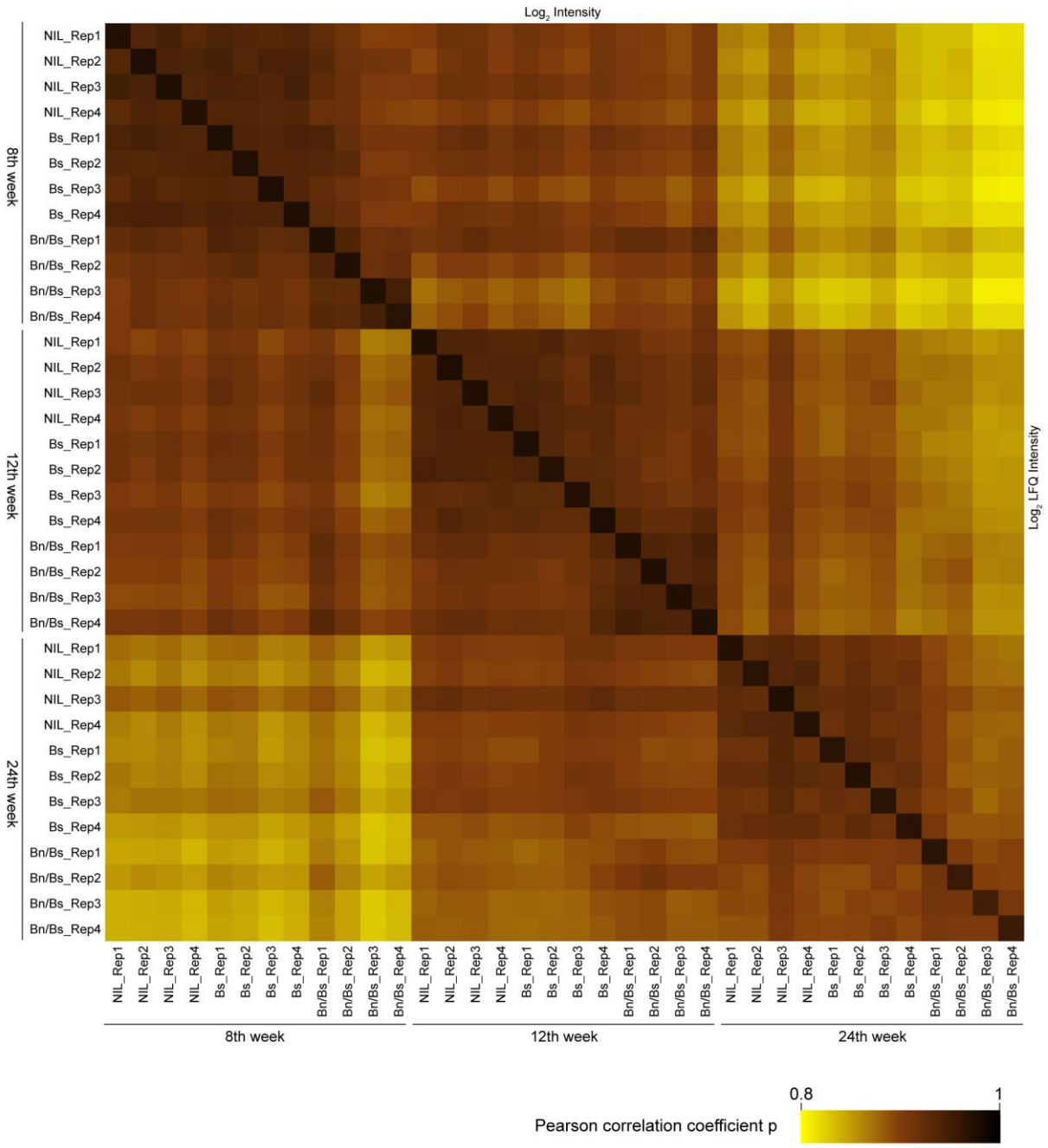
Title

Comparative proteome and metabolome analyses of latex-exuding and non-exuding *Taraxacum koksaghyz* roots provide insights into laticifer biology

Authors

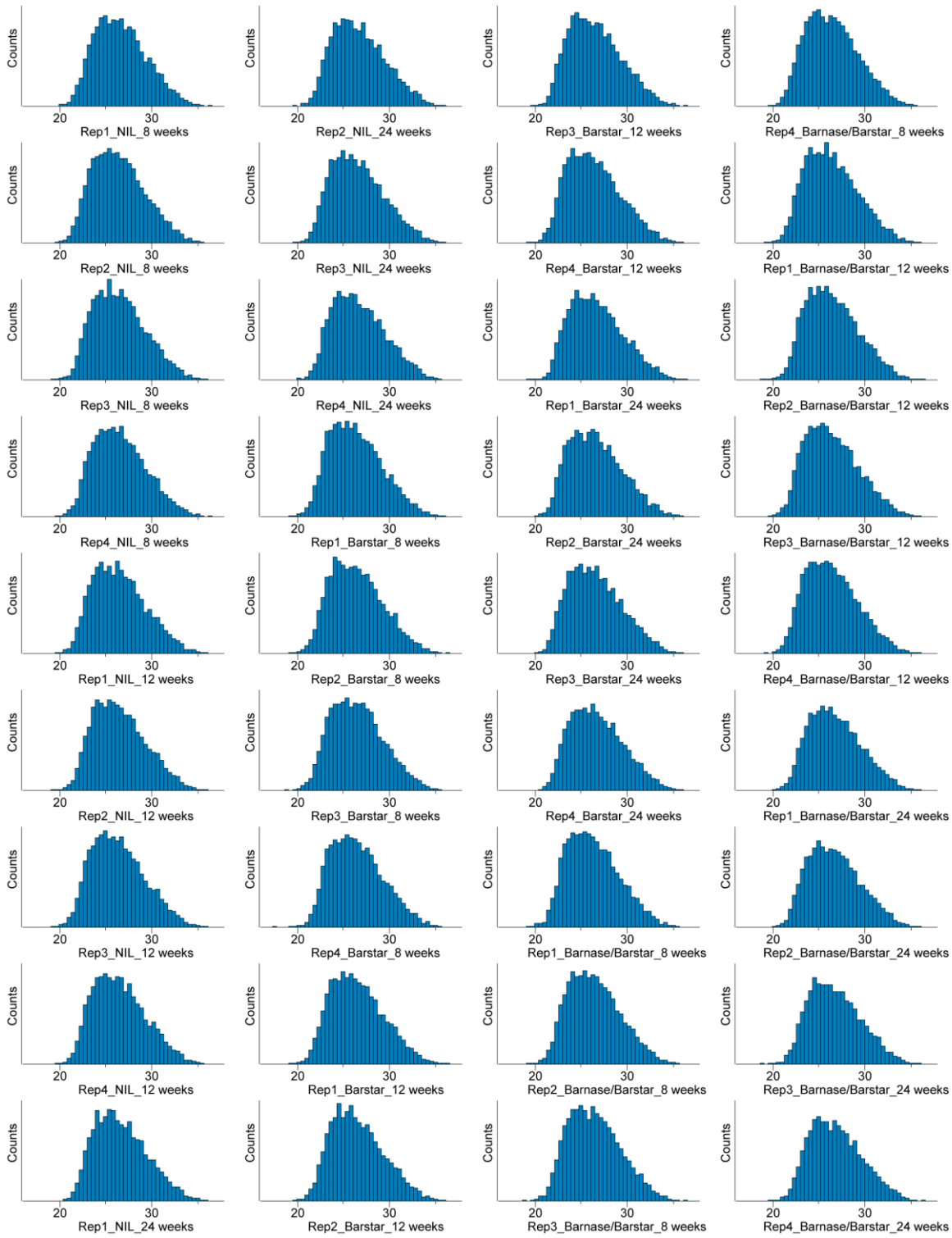
Vincent Alexander Benninghaus[1], Nicole van Deenen[2], Boje Müller[1], Kai-Uwe Roelfs[1], Ines Lassowskat[2], Iris Finkemeier[2], Dirk Prüfer[1,2] and Christian Schulze Gronover[1]

[1]Fraunhofer Institute for Molecular Biology and Applied Ecology IME, Schlossplatz 8, Muenster 48143, Germany. [2]Institute of Plant Biology and Biotechnology, University of Muenster, Schlossplatz 7-8, Muenster 48143, Germany. E-mail: vincent.benninghaus@ime.fraunhofer.de, nicole.vandeenen@uni-muenster.de, boje.mueller@ime.fraunhofer.de, roelfs@uni-muenster.de, ilasso@uni-muenster.de, iris.finkemeier@uni-muenster.de, dpruefer@uni-muenster.de, christian.schulze.gronover@ime.fraunhofer.de



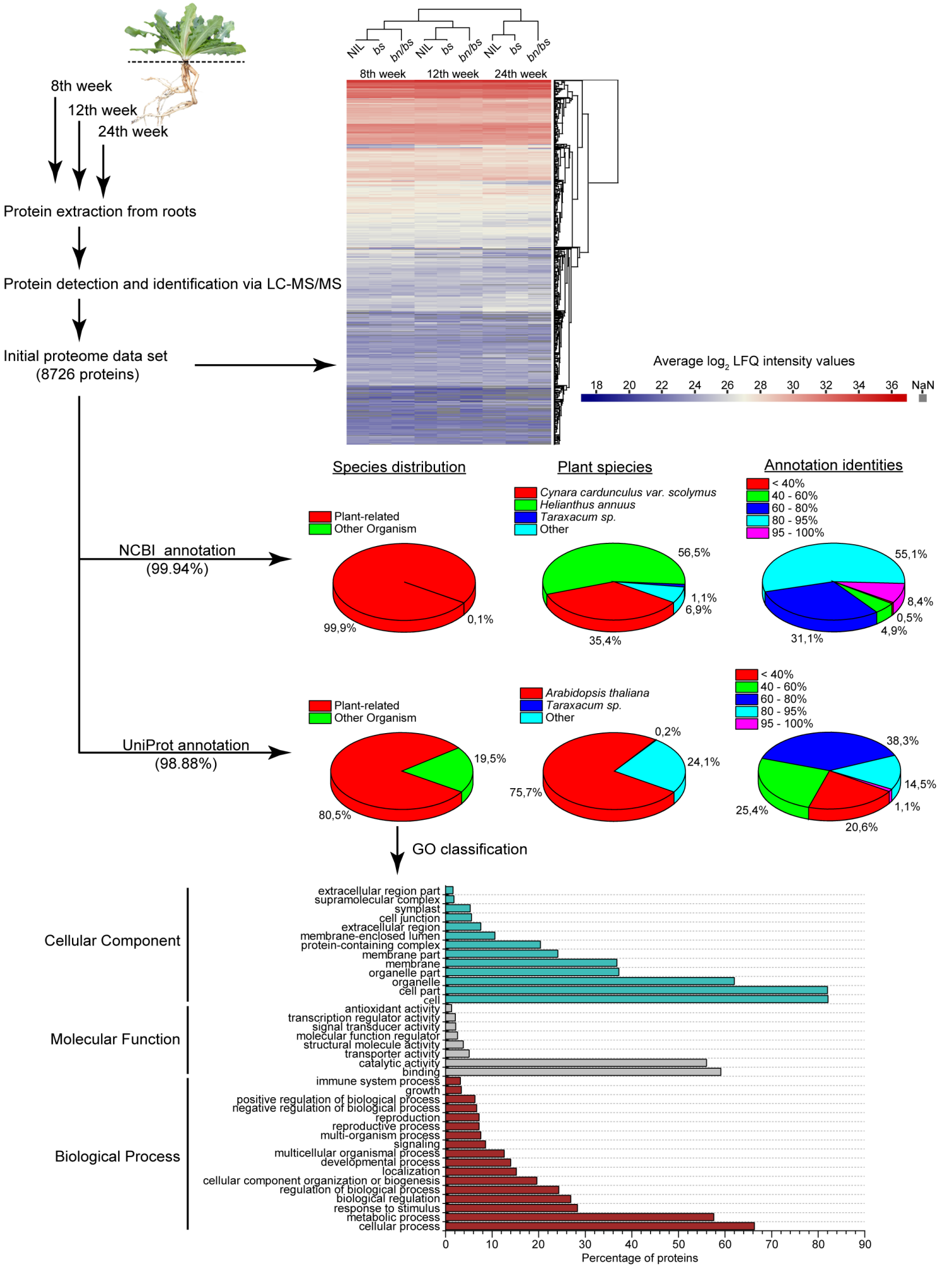
Supplementary Figure 1: Validation of the proteome dataset normalization procedure using MaxQuant.

Whole-root protein extracts from 8-, 12- and 24-week-old *T. koksaghyz* plants of all three genotypes (Near isogenic line, NIL; barstar, *bs*; barnase, *bn/bs*) were used for the proteome analysis. The tryptic digested peptides were analysed using an ultra-high performance liquid chromatography-system coupled to a quadrupole orbitrap exactive high-performance mass spectrometer. The resulting raw data of all 36 samples (4 replicates of each genotype per time point) were processed using the MaxQuant software. To confirm the reliability of the automatic normalization procedure of MaxQuant, the obtained $\log_2$ raw intensities were plotted against the calculated $\log_2$ LFQ intensities. The color range represents the corresponding different values of the Pearson correlation coefficient (ρ) from 0.8 (yellow) to 1 (black).



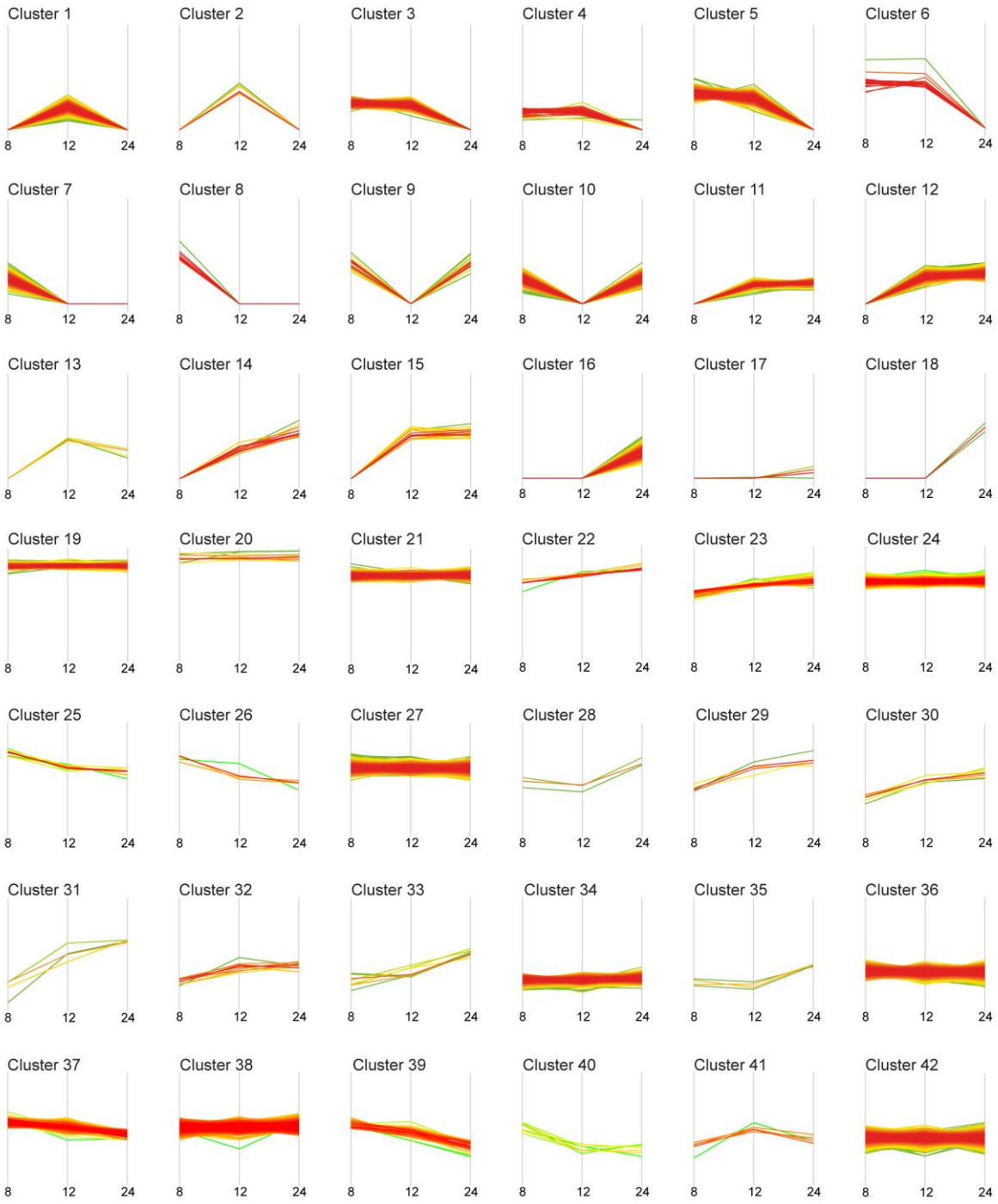
Supplementary Figure 2: Histograms of the $\log_2$ LFQ intensities of all 36 proteome samples.

Whole-root protein extracts from 8-, 12- and 24-week-old *T. koksaghyz* plants of all three genotypes (Near isogenic line, NIL; barstar, *bs*; barnase, *bn/bs*) were used for the proteome analysis. At each point in time, 4 replicates (Rep1-4) of each genotype were measured. The extracted proteins were subjected to a trypsin digestion and subsequently analysed using an ultra-high performance liquid chromatography-system coupled to a quadrupole orbitrap exactive high-performance mass spectrometer. The obtained raw data were processed using MaxQuant software. The calculated $\log_2$ LFQ intensities were depicted as histograms, to verify and guarantee the normal distribution of all 36 samples.


8th week
12th week
24th week
Protein extraction from roots
Protein detection and identification via LC-MS/MS
Initial proteome data set
(8726 proteins)
NIL
bs
bn/bs
8th week
12th week
24th week
Average log2 LFQ intensity values
18 20 22 24 26 28 30 32 34 36 NaN
NCBI annotation
(99.94%)
Species distribution
Plant-related
Other Organism
99,9%
0,1%
Plant spiecies
Cynara cardunculus var. scolymus
Helianthus annuus
Taraxacum sp.
Other
56,5%
1,1%
6,9%
35,4%
Annotation identities
< 40%
40 - 60%
60 - 80%
80 - 95%
95 - 100%
55,1%
8,4%
0,5%
4,9%
31,1%
UniProt annotation
(98.88%)
Plant-related
Other Organism
19,5%
80,5%
Arabidopsis thaliana
Taraxacum sp.
Other
0,2%
24,1%
75,7%
< 40%
40 - 60%
60 - 80%
80 - 95%
95 - 100%
38,3%
14,5%
1,1%
20,6%
25,4%
GO classification
Cellular Component
Molecular Function
Biological Process
extracellular region part
supramolecular complex
symplast
cell junction
extracellular region
membrane-enclosed lumen
protein-containing complex
membrane part
membrane
organelle part
organelle
cell part
cell
antioxidant activity
transcription regulator activity
signal transducer activity
molecular function regulator
structural molecule activity
transporter activity
catalytic activity
binding
immune system process
growth
positive regulation of biological process
negative regulation of biological process
reproduction
reproductive process
multi-organism process
signaling
multicellular organismal process
developmental process
localization
cellular component organization or biogenesis
regulation of biological process
biological regulation
response to stimulus
metabolic process
cellular process
0 10 20 30 40 50 60 70 80 90
Percentage of proteins

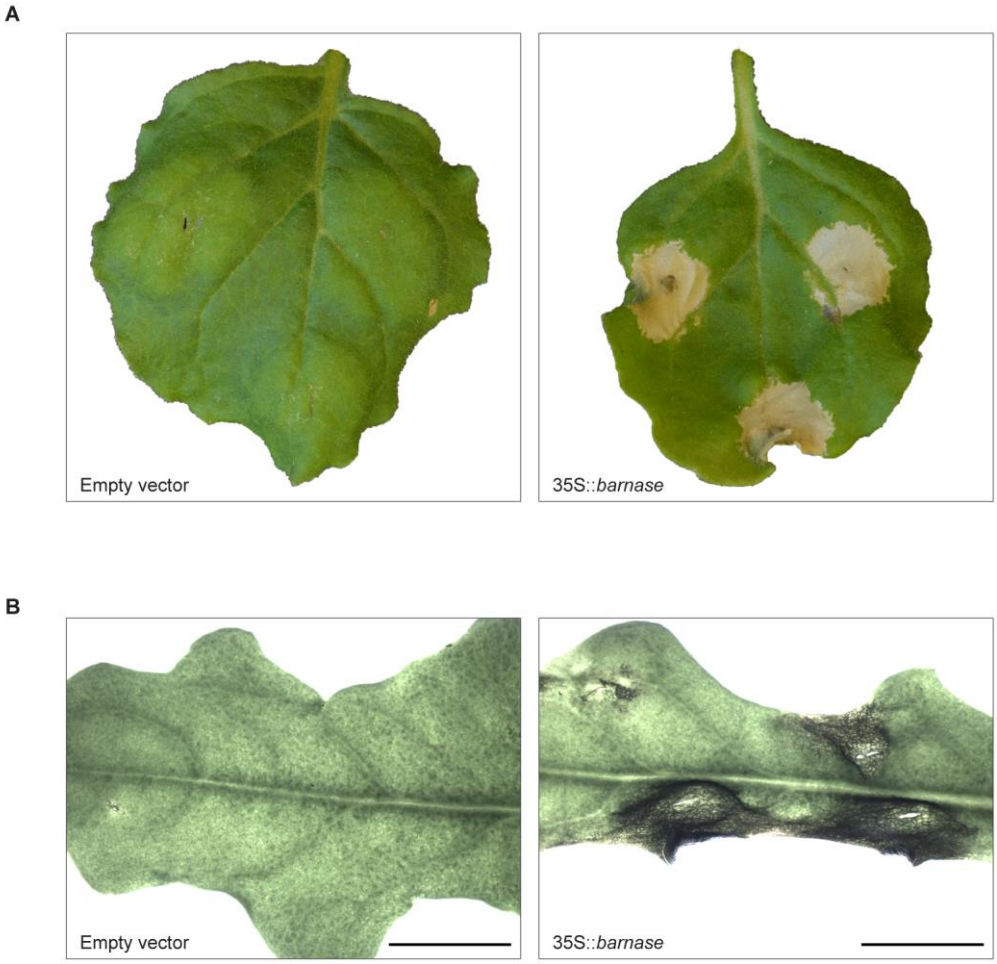
Supplementary Figure 3: Workflow used to generate the initial proteome dataset.

The root proteome analysis was performed from 8-, 12- and 24-week-old *T. koksaghyz* plants of all three genotypes (Near isogenic line, NIL; barstar, *bs*; barnase, *bn/bs*). At each point in time, 4 replicates of each genotype were chosen. After analysing the tryptic digested peptides using an ultra-high performance liquid chromatography-system coupled to a quadrupole orbitrap exactive high-performance mass spectrometer, the obtained raw data were processed using MaxQuant and searched against the dandelion database (Genome Warehouse, http://bigd.big.ac.cn/gwh/, accession number PRJCA000437). The resulting data were then further processed and quality filtered, whereby an initial proteome matrix encompassing 8726 protein groups was generated. For hierarchical cluster analysis, the average $\log_2$ LFQ intensities of each genotype per time point were calculated and compared. Pie charts represent information about protein annotation results. Gene ontology classification of the initial proteome dataset was performed by plotting the corresponding GO annotation.



Supplementary Figure 4: Temporal protein accumulation patterns of the near-isogenic line depicted as 42 clusters.

Depicted are the 42 protein clusters that were calculated out of the heatmap (Fig. 3, Supplementary Table 5). Assembled are those proteins that had a similar temporal expression pattern from the 8th to the 24th week. The higher the lines in the diagram, the higher the protein intensities and vice versa.

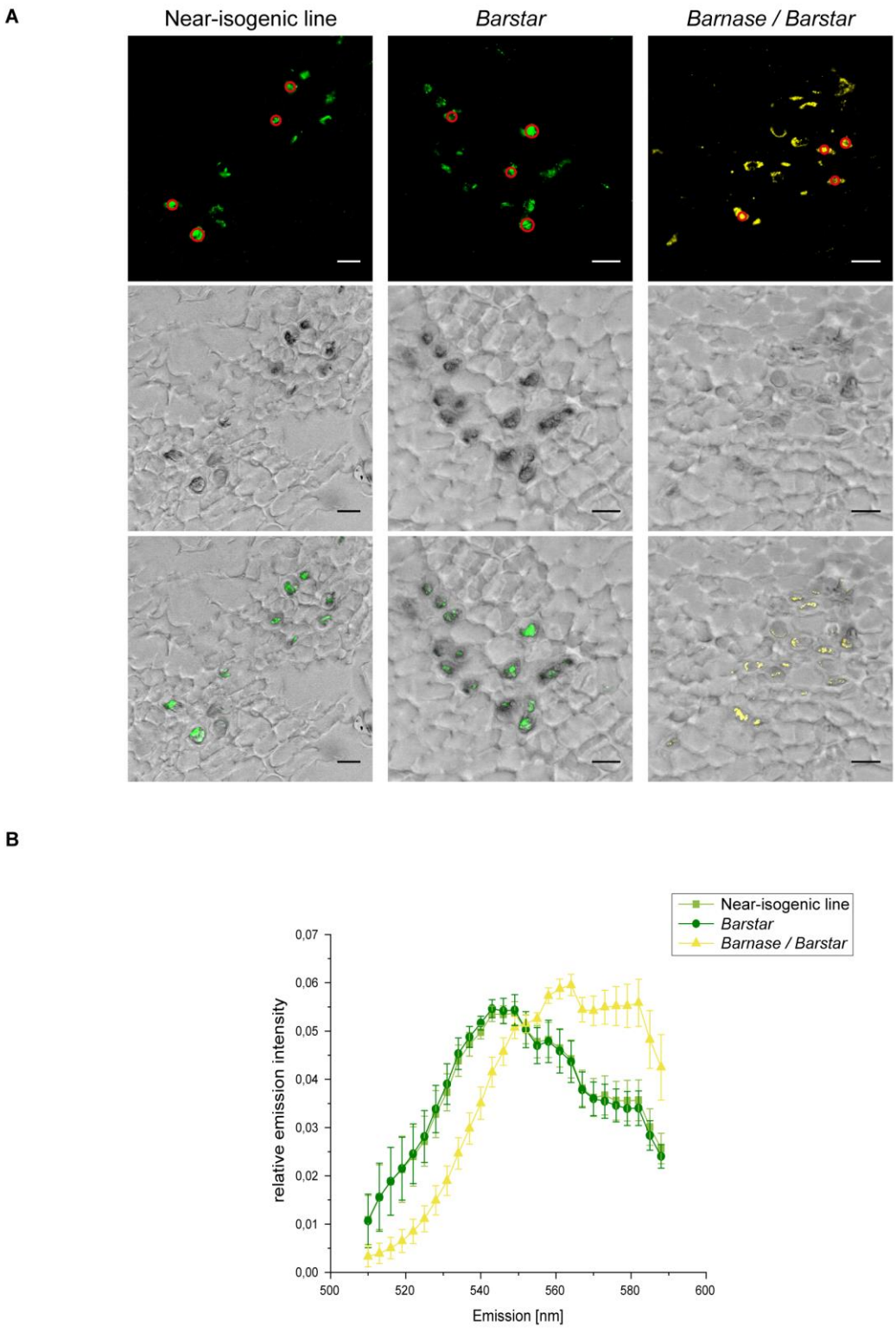


Supplementary Figure 5: Examination of the cytotoxic activity of the generated barnase.

(A,B) Agro-infiltration assays in (A) *Nicotiana benthamiana* and (B) *Taraxacum koksaghyz* leaves for evaluation of barnase enzymatic activity. Vector controls (left), *barnase* gene under the control of the *Cauliflower mosaic virus* 35S promoter (right). Scale bars: 1 cm.

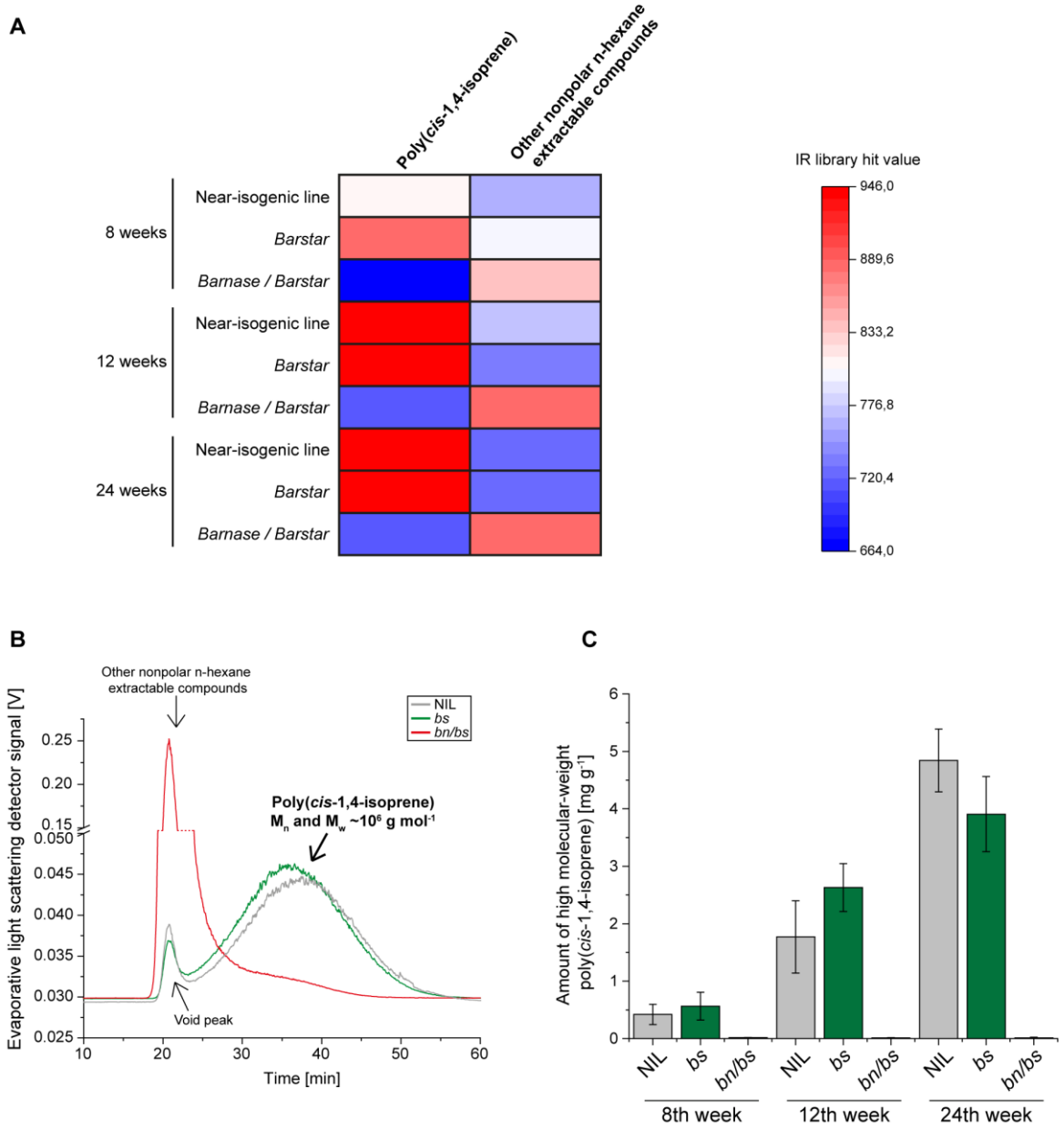
Supplementary Figure 6: Plant morphology over time of the different genotypes.

Depicted are the latex-exuding control (Near isogenic line, NIL; barstar, *bs*) and non-exuding (barnase, *bn/bs*) plants at three different time points. The corresponding root crown cuttings are shown at the bottom right of each picture. Scale bar: 3.6 cm.



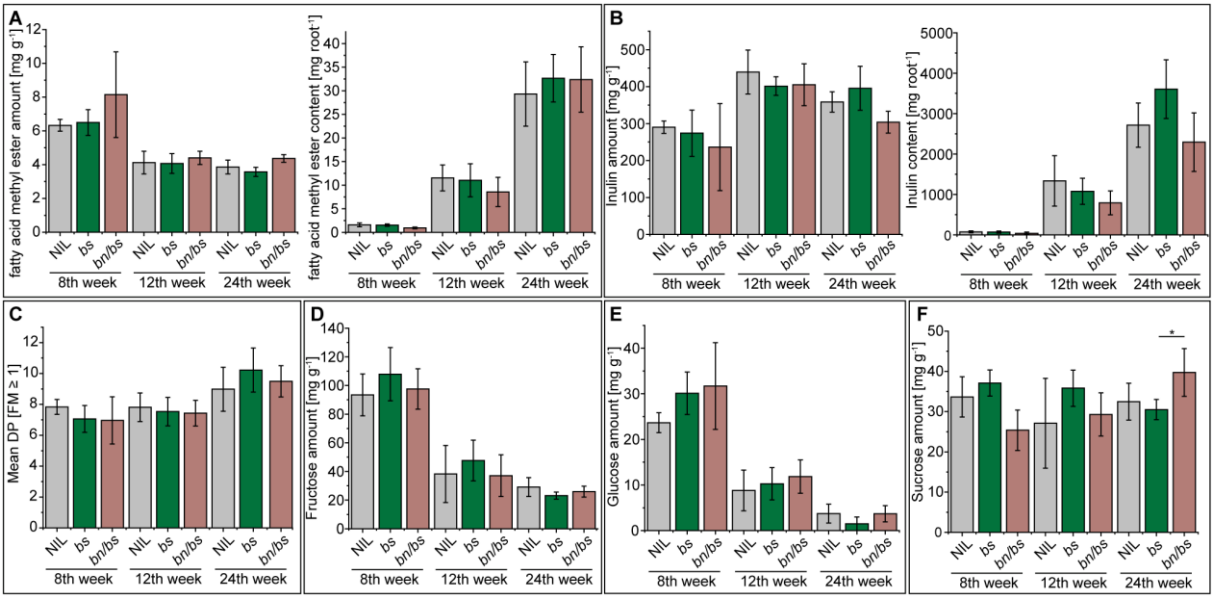
Supplementary Figure 7: CLSM analysis of cross-sections from the three genotypes.

A confocal laser scanning microscope (CLSM) was used for recording the microscopic images. (A) Cryosections (20 µm) were cut from fresh root material of latex-exuding control (Near isogenic line, NIL; barstar, *bs*) and non-exuding (barnase, *bn/bs*) plants, and stained with the lipophilic fluorescent dye Nile red. Red circles show ROI positions. Scale bars: 20 µm. (B) Determination of the Nile red emission spectra performed with lambda scans (Excitation: 488 nm, Emission wavelength range: 510-590 nm) revealed a shift towards longer wavelength in non-exuding (*bn/bs*) plants, triggered by the solvatochromaticity of Nile red (Diaz et al., 2008).



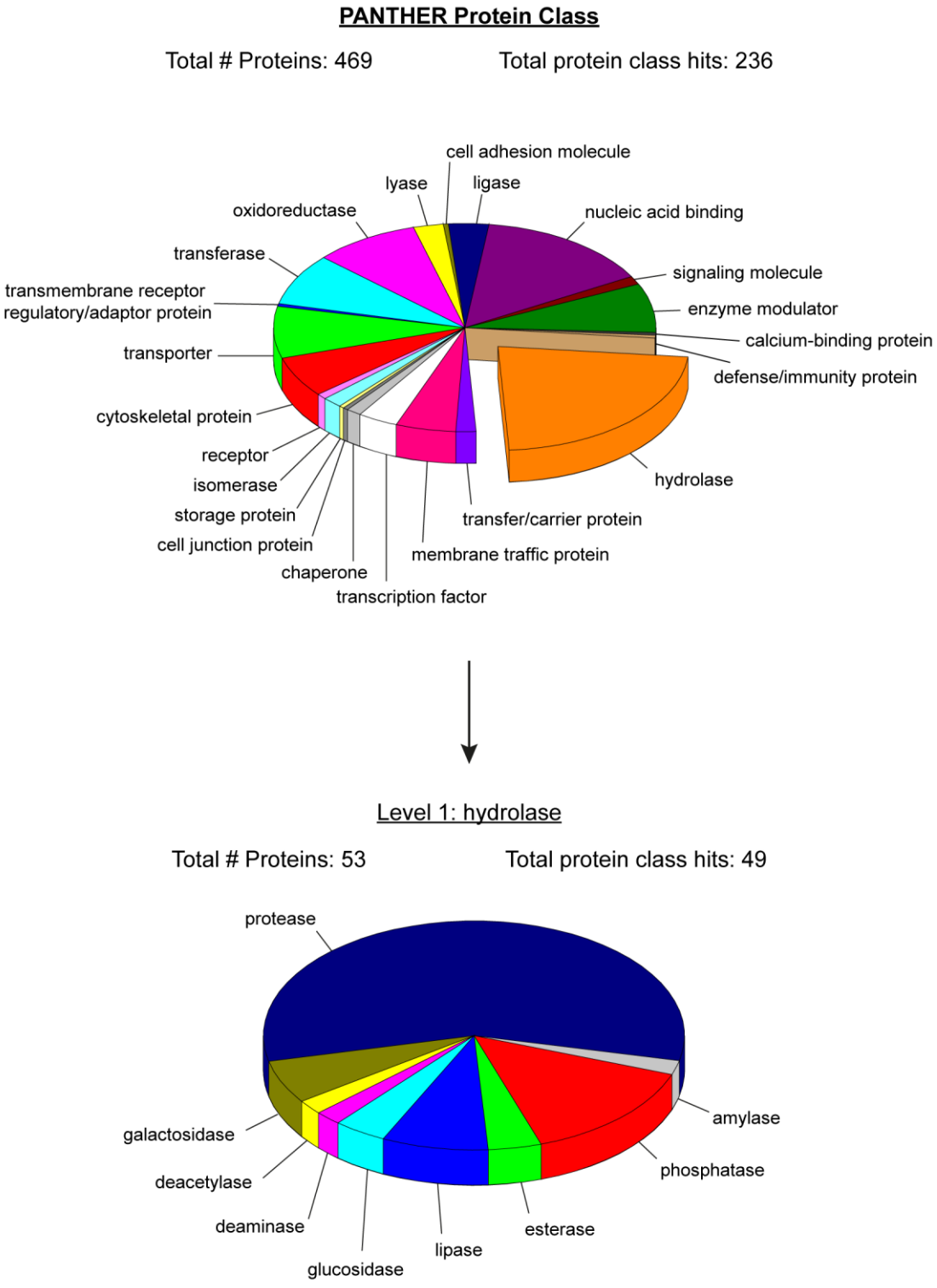
Supplementary Figure 8: Content of high molecular-weight poly(*cis*-1,4-isoprene) in the roots.

Powdered root samples from 8-, 12- and 24-week-old *T. koksaghyz* plants of all three genotypes (Near isogenic line, NIL; barstar, *bs*; barnase, *bn/bs*) were subjected to accelerated solvent extraction. (A) The obtained hexane root extracts were measured by Fourier-Transform Infrared Spectroscopy and the resulting values of the best library hits were recorded and depicted as headmap. (B,C) The extracts were dissolved with toluene at 5 mg ml^{-1} and were then fractionated by Thermal Field-Flow Fractionation. (B) The illustrated graph exemplifies the signal detected by the evaporative light scattering detector from 24-week-old *T. koksaghyz* plants of all three genotypes. Here, the extracts obtained from the control plants NIL and *bs* contained almost exclusively high molecular-weight poly(*cis*-1,4-isoprene) with a number average molecular mass (M_n) and a mass average molar mass (M_w) of about 10^6 g mol^{-1}. In contrast, only negligible amounts of high molecular-weight poly(*cis*-1,4-isoprene) were present in the roots of non-exuding (*bn/bs*)-plants. This led to a change in the solubility equilibrium, so that predominantly other nonpolar hexane-soluble compounds were extracted and subsequently measured. (C) Concentrations of high molecular-weight poly(*cis*-1,4-isoprene) calculated from the MALS signal.

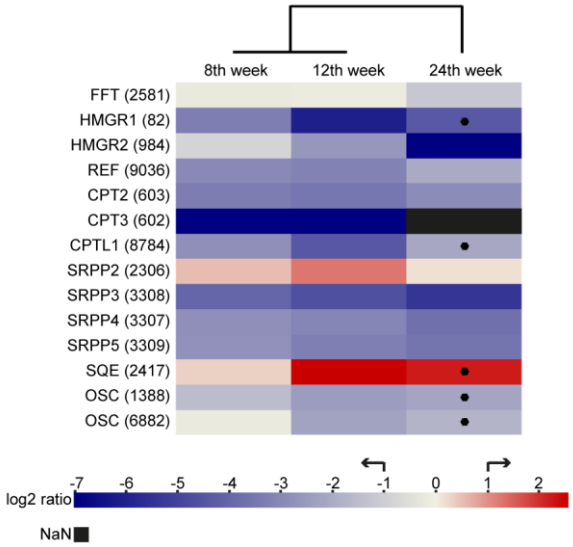


Supplementary Figure 9: Primary metabolites in the roots.

Roots were sampled at 8, 12, and 24 weeks old. Data are means (±SD): for the 8th week five pooled samples of four plants each were sampled, and for the 12th and 24th weeks six individual plants were sampled. (A,B) Metabolite concentration and the total content per root were monitored for fatty acid methyl esters and inulin. (C-F) The mean degree of polymerization of inulin and metabolite concentrations of fructose, glucose and sucrose were also calculated. The fatty acid methyl ester amount and content were calculated based on the detection of palmitic acid (16:0), oleic acid (18:1) and linoleic acid (18:2). Due to non-normal distribution of the data, significance levels among the genotypes at each time-point were assessed using non-parametric Mann–Whitney U-tests (P<0.01).



Supplementary Figure 10: Protein class annotation of less-abundant proteins in the *bn/bs* plants. Protein class annotation was performed from all proteins detected as decreased differentially accumulated (Fig. 5) using the online tool PANTHER (Protein ANalysis THrough Evolutionary Relationships) classification system.



Supplementary Figure 11: Differentially accumulated proteins within inulin and terpene-backbone metabolism.

Proteins that showed a $\log_2$-ratio of > 1 or < -1 and *P*<0.05 between latex-exuding (barstar, *bs*) and non-exuding (barnase, *bn/bs*) plants were considered as differentially accumulated; in addition, proteins that were not detectable (invalid $\log_2$ ratio) or that were almost absent in one of the genotypes (quantified only for one out of four replicates, thus not allowing for the generation of a *P*-value) were also taken into account; however, a prerequisite for this was the detection of these proteins in the other comparative genotype in at least three out of four replicates. Excluded from this are the differentially accumulated proteins (DAPs) marked with a polygon, since they did not show a significant *P*<0.05 or were prior determined as DAPs between both controls (near-isogenic line, NIL and *bs*), and thus were excluded in this evaluation. Non assigned number, NaN; fructan: fructan 1-fructosyltransferase, FFT; 3-hydroxy-3-methylglutaryl-coenzyme A reductase, HMGR; rubber elongation factor, REF; *cis*-prenyltransferase, CPT; *cis*-prenyltransferase like, CPTL1; small rubber particle protein, SRPP; squalene epoxidase, SQE; oxidosqualene cyclase, OSC.